Patients

Newly diagnosed patients with T1, N1-3, M0 or T2-4, N0-3, M0 breast cancers (with cancer ≥ 1.5 cm), negative for estrogen and progesterone receptors, defined as < 1% nuclear staining by immunohistochemistry, and HER2/neu 0 or 1+ by immunohistochemistry, or HER2 non-amplified by FISH, were eligible for this trial. Patients could not have had previous chemotherapy.

Study Design and Treatment Plan

The clinical protocol was approved by the Dana-Farber/Harvard Cancer Center IRB. After written informed consent, blood was taken for research analyses, including germline BRCA1 genotyping for investigational purposes only. A core biopsy was performed to obtain tumor tissue for study, and a radio-opaque clip was placed in the tumor bed. For patients with a clinically negative axilla, sentinel lymph node biopsy was required before therapy; for patients with a clinically positive axilla, either a sentinel node biopsy or a needle aspiration was performed. If the needle aspiration showed malignant cells, no sentinel lymph node biopsy was required. Patients received 4 treatments of cisplatin at 75mg/m^2 every 21 days. Patients then received definitive surgery, including an axillary lymph node dissection in patients with positive fine needle aspirate or sentinel lymph node biopsy. The pathology specimen was evaluated for chemotherapy response, with focused sampling of the tumor bed marked by the radio-opaque clip. The protocol stipulated that standard adjuvant therapy should

then be administered, with the regimen left to the discretion of the treating physician.

Sample Size Justification and Statistical Methods

The protocol document specified that the primary endpoint would be clinical response rate (clinical complete or partial response using RECIST criteria; the responses were mostly based on MRI but sometimes based on palpation). The study had a two stage design: if fewer than 4 of the first 12 patients treated had a clinical response then accrual of the study was to close; if 4 or more treated had a clinical response then an additional 15 patients were to be entered (for a total of 27 treated patients); and if at least 10 patients out of the 27 had a clinical response then the cisplatin treatment was to be deemed worthy of further study in triple negative patients. (This design was stated to have an alpha level of 0.025 if the true clinical response rate is 20% and a power of 98% if the true clinical response rate is 60%, with calculations based on exact binomial distributions). The protocol also had a stopping rule based on pCRs: if none were observed in the first 12 patients then accrual would end, and the protocol implied that 4 pCRs out of the 27 patients entered would be considered acceptable. For both overall clinical response and pCR, the protocol specified that conditional 95% confidence intervals (CIs) were to be used because of the two stage design (using the method of Atkinson and Brown[26]). The protocol had no stopping rule based on major pathologic response (Miller-Payne 3, 4, or 5) and hence had no rule for calculating a CI for this endpoint. In order to be

consistent with methods used for the other outcome measures, we assumed a two stage design for major pathologic response similar to that for clinical response and used a conditional CI (which turned out to be 1% larger than the unconditional CI).

Fisher exact tests were used to evaluate if the three-valued variable p53 or the binary variables (axillary lymph node positivity, BRCA1 methylation, p63/p73>2.0, or any of the binary variables based on p53) were associated with any of the response outcome measures. Continuous variables were divided into quartiles and an ordered Fisher exact test (equivalent to an exact Wilcoxon rank sum test) was used to test if any of these variables were associated with response. Step up logistic regressions were used to explore the association of combinations of variables and response. The Wilcoxon rank sum test was also used to evaluate whether BRCA1 expression levels were different in tumors with and without BRCA1 promoter methylation. The Spearman rank correlation was used to evaluate the relationship of ID4 and BRCA1 mRNA. All P values were two-sided. Unless otherwise stated, P values are not adjusted for multiple comparisons. The plan was to use Holm's method of adjustment for multiple comparisons for the non-gene array data. Since only one variable (age) was significant after this adjustment, the Holm's method is equivalent to the Bonferroni method.

Microarray-based gene expression measurements and gene signatures were tested for association with response using the Pearson correlation test; P-values are two-sided and are uncorrected unless stated otherwise.

Specimen Analysis

Histologic sections of all pre and post-treatment specimens were reviewed by a single pathologist (A.R.), using the Miller-Payne scale[25] to quantify pathological response to treatment. The Miller-Payne (MP) scale is based on the estimated percent reduction in invasive tumor volume and cellularity in the breast: MP1, no reduction in tumor; MP2, modest reduction up to 30%; MP3, reduction of 30-90%; MP4, >90% reduction but with scant residual invasive disease; MP5, complete response with no residual invasive disease. The MP score has been shown to correlate with both overall and disease-free survival.[25] The Miller-Payne response measure was modified to include evaluation of residual disease in the axillary nodes, in addition to the breast, to determine the response score. BRCA1 genotyping for research purposes was performed by the Exon Grouping Analysis (EGAN) method on germline DNA extracted from peripheral blood leukocytes.

EGAN Genotyping

EGAN is based on Conformation Specific Gel Electrophoresis (CSGE)[25,26]. All coding exons and surrounding intronic sequences were amplified by PCR and analyzed on ABI-377 instruments. PCR fragments with aberrant mobility were sequenced. This method has been compared directly to standard sequencing using a blinded patient set and has shown a sensitivity of 97.4% in detecting *BRCA1* and *BRCA2* sequence changes (A. Miron, manuscript in preparation).

Array Analysis

For gene expression array analysis, two 5µm frozen sections were stained with hematoxylin and eosin, manually scraped to remove normal stroma and enrich for tumor cells, and RNA was extracted using the Absolutely RNA Nanoprep kit (Stratagene). First round linear amplification was performed using RiboAmp HS kit (Arcturus), followed by in vitro transcription using the Affymetrix IVT kit. Gene expression profiling was then performed on Affymetrix U133 Plus 2.0 microarrays at the Dana-Farber/Harvard Cancer Center array core facility. The complete gene expression array dataset will be available on the NCBI GEO database (accession submission in progress).

Quantitative RT-PCR

For quantitative RT-PCR, tumor cells were microdissected from two 7µm frozen sections and RNA was extracted using the Micro RNA Isolation kit (Stratagene). Six samples did not have adequate material for qRT-PCR analysis. cDNA was synthesized from total RNA using random hexamer primers and the SuperScript II system for RT-PCR (Invitrogen). Q-PCR analysis was carried out using iQ™ SYBR Green mix (Bio-Rad). The RPLP0 transcript was used as an internal control for normalization of relative expression levels. Primers used for qPCR were designed across exons to avoid amplification of genomic DNA. Oligo sequences for BRCA1 cDNA primers: E1F (AGGAGGCCTTCACCCTCTGCTCT), E2R (TTCAACGCGAAGAGCAGATAAATCCA), E16F (TCAACAAAAGAATGTCCATGGTGGTGTC), E17R (GTGATGTGGTGTTTTCTGGCAAACTTGT), E19F (GGGTGACCCAGTCTATTAAAGAAAGAAAAATGCT) and E20R (TTCTTCCATTGACCACATCTCCTCTGACTT). Primers used for the RPLP0 housekeeping gene were: RPLP0F (ATCAACGGGTACAAACGAGTCCTG) and RPLP0R (AAGGCAGATGGATCAGCCAAGAAG). Primer sequences for ΔNp63 were: ΔNp63F (GGAAAACAATGCCCAGACTC) and ΔNp63R (GTGGAATACGTCCAGGTGGC). Primer sequences for TAp73: TAp73F (GCACCACGTTTGAGCACCTCT) and TAp73R (GCAGATTGAACTGGGCCATGA). Sequencing of the p53 coding region was performed on cDNA from microdissected tumor cells using the following primers: p53-1F (CAAGCAATGGATGATTTGATG), p53-1R (CTTCTTTGGCTGGGGAGAG), p53-2F (CACATGACGGAGGTTGTGAGG) and

p53-2R (TTTTTATGGCGGGAGGTAGA). BRCA1 promoter methylation was performed by methylation-specific PCR as in [27].

Tumor subtype determination

To determine tumor subtype, a reference set of 61 primary breast tumor samples (52 independent tumors and 9 array replicates) were prepared using the same methodology as the trial samples, including an initial round of linear amplification, and expression profiling on the same microarray platform. Normalized expression values were calculated from raw data using the RMA algorithm. Reference samples and trial samples were processed together. The 552 "intrinsic" genes[19]were annotated by clone ID, which we mapped to 473 unique UniGene identifiers using the SOURCE database (http://smd.stanford.edu/cgi-bin/source/sourceSearch). Using annotation data from Bioconductor, 388 of these Unigene identifiers were mapped to 1032 probe sets on our microarray platform. For each of the 388 “intrinsic” genes, we eliminated redundant measurements by retaining only a single probe set with the highest variance of log2 intensity over all samples.

The expression value of each gene was centered by subtracting its mean over the 61 reference samples. To assign subtypes to the reference set we first clustered the samples in the reference set alone (supplemental Figure 1). We identified the five most distinct clusters, each of which we assigned to a subtype by choosing the highest Pearson correlation coefficient between the cluster centroid and the centroids of Sorlie et al.[19] To assign subtypes to the trial samples, we clustered the trial samples together with the reference samples, again centering the genes using only the reference samples. All clustering was performed using 1 - Pearson correlation distance and average linkage. The R scripts used to perform this analysis follow on the next page.

.

Clustering Analysis For Cisplatin Trials

November 17, 2008

1 Clustering Based on Intrinsic Genes

load platinum dataset, processed by RMA and dChip (implemented in R, without log transformation).

```
> library(affy)
> library(squash)
> load("data/platinum.rma.RData")
```

Load intrinsic genesets (Sorlie et al. PNAS July 8, 2003)

```
> load("data/Intrinsic_cent_pnas.RData")
> dim(Centroids)
```

```
[1] 552   5
```

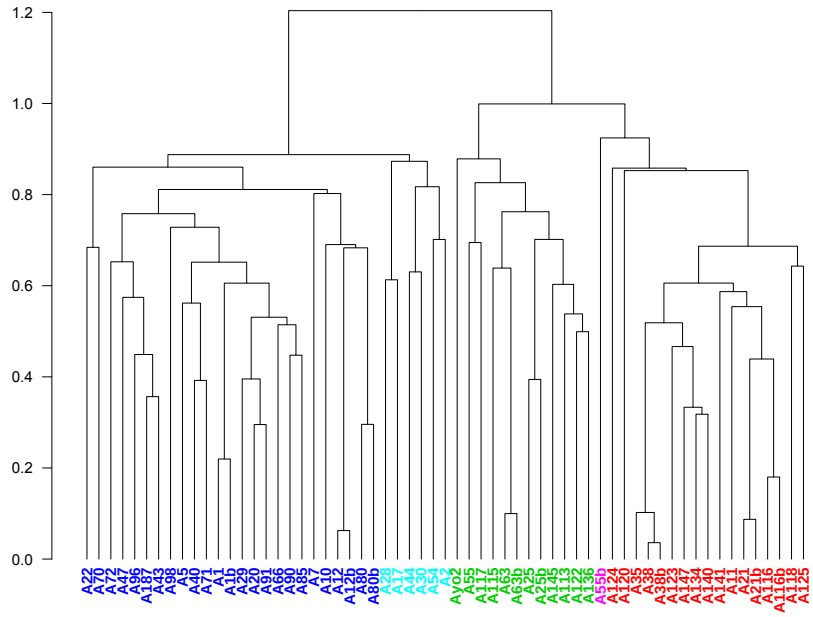
Now we retrieve the expression profile of these intrinsic genes from platinum dataset, using non-cisplatin treated sample only. In case multiple Affymetrix probesets mapped to one intrinsic gene, the probeset with the largest cross-sample variance are selected to represent the gene.

```
> profile <- exprs(platinum.rma)
> rownames(profile) <- mget(rownames(profile), hgu133plus2UNIGENE, ifnotfound = NA)
> profile <- profile[rownames(profile) %in% rownames(Centroids), ]
> profile <- profile[order(apply(profile, 1, function(x) var(x, na.rm = T)), decreasing =
+     ]
> profile <- profile[!duplicated(rownames(profile)), ]
> outliers <- c("AR2006110811.CEL", "AR2006110825.CEL", "AR2007010390.CEL", "AR2007010391.
> profile <- profile[, which(!colnames(profile) %in% outliers)]
> indx <- which(!pData(platinum.rma)[colnames(profile), 15])
> profile <- profile[, indx]
> colnames(profile) <- pData(platinum.rma)[colnames(profile), 2]
> dim(profile)
```

```
[1] 388  61
```

First we retried expression profile of 388 intrinsic genes (about 70%) in 61 samples from platinum dataset. The expression profile of each gene wass centered by the mean of all samples. Hierarchical clustering was performed based Pearson's correlation matrix and average linkage method, as described in the original study.

```
> profile <- t(apply(profile, 1, function(x) x - mean(x)))
> C.ref <- hclust(cor.dist(t(profile), method = "pearson"), method = "average")
> dhc <- as.dendrogram(C.ref)
> colLab <<- function(n) {
+     if (is.leaf(n)) {
+         a <- attributes(n)
+         i <<- i + 1
+         attr(n, "nodePar") <- c(a$nodePar, list(lab.col = mycols[attr(n, "label")],
+             lab.cex = 1, pch = ""))
+     }
+     n
+ }
> mycols <- sapply(cutree(C.ref, k = 5), function(x) switch(x, 4, 3, 2, 5, 6))
> j <- sapply(colnames(profile), function(x) which(pData(platinum.rma)[, 2] == x))
> mycols[colnames(profile)[which(pData(platinum.rma)[j, 13] > 2)]] <- "#FDD017"
> mycols[colnames(profile)[which(pData(platinum.rma)[j, 13] < 2)]] <- 1
> i <- 0
> dhc <- dendrapply(dhc, colLab)
> par(las = 1, cex = 1.2, font = 2, mar = c(6, 6, 4, 2) + 0.1)
> plot(dhc, edge.root = F)
```



We compute the centroids for each of the five clusters in previous analysis and compared them to the original intrinsic centroids reported by Sorlie et al. Each cluster was assigned to a certain intrinsic subtype to which it indicated the highest correlation. However for cluster 4, its centroid correlated with "Lum A" with a coefficient of 0.369, which was the highest among

the 5. But since cluster 1 has a much higher correlation to "Lum A" centroids, we assigned this cluster to "Lum B".

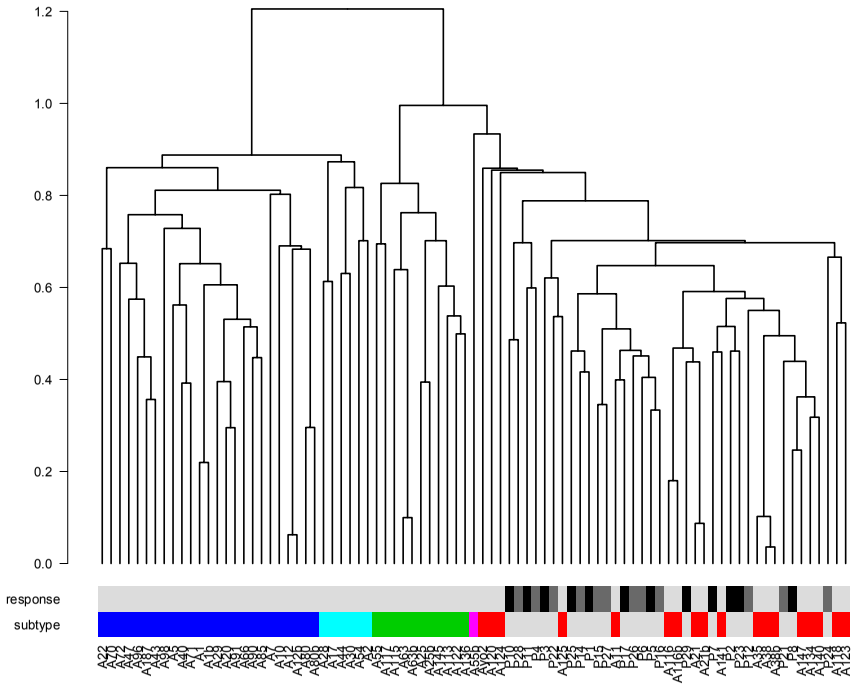
```
> plat.cent <- sapply(1:4, function(x) apply(profile[, which(cutree(C.ref, k = 5) ==
+     x)], 1, mean))
> plat.cent <- cbind(plat.cent, profile[, which(cutree(C.ref, k = 5) == 5)])
> x <- apply(plat.cent, 2, function(x) cor(x, Centroids[rownames(plat.cent), ]))
> rownames(x) <- colnames(Centroids)
> x

            [,1]        [,2]        [,3]        [,4]        [,5]
Lum A  0.6380363 -0.32876483 -0.58017200  0.36874239 -0.03850577
Lum B -0.1945348  0.29846881  0.05814220  0.01613455 -0.10129076
ERBB2 -0.3224728  0.40150529  0.14587610 -0.08970220  0.02086657
Basal -0.6573541  0.03449855  0.76574421 -0.45979608  0.10951468
Norm   0.1057813 -0.21191810  0.01351511 -0.17130957  0.20104769

> plat.clust <- c("Lum A", "ERBB2", "Basal", "Lum B", "Norm")
```

Now we redo hierarchical clustering in the pooled dataset with both reference set and cisplatin trials. Note the gene are still centered by mean of reference set.

```
> profile <- exprs(platinum.rma)
> rownames(profile) <- mget(rownames(profile), hgu133plus2UNIGENE, ifnotfound = NA)
> profile <- profile[rownames(profile) %in% rownames(Centroids), ]
> profile <- profile[order(apply(profile, 1, function(x) var(x, na.rm = T)), decreasing =
+     ]
> profile <- profile[!duplicated(rownames(profile)), ]
> outliers <- c("AR2006110811.CEL", "AR2006110825.CEL", "AR2007010390.CEL", "AR2007010391.
> profile <- profile[, which(!colnames(profile) %in% outliers)]
> indx <- which(!pData(platinum.rma)[colnames(profile), 15])
> colnames(profile) <- pData(platinum.rma)[colnames(profile), 2]
> profile <- t(apply(profile, 1, function(x) x - mean(x[indx])))
> C <- hclust(cor.dist(t(profile), method = "pearson"), method = "average")
> dhc <- as.dendrogram(C)
> j <- sapply(colnames(profile), function(x) which(pData(platinum.rma)[, 2] == x))
> type.ref <- plat.clust[cutree(C.ref, k = 5)]
> names(type.ref) <- names(cutree(C.ref, k = 5))
> mat <- data.frame(subtype = as.factor(type.ref[names(j)][colnames(profile)]), response =
+     13] > 2, "#000000", "#696969"))
> rownames(mat) <- names(j)
> mat[, 1] <- sapply(cutree(C, k = 5), function(x) switch(x, 4, 3, 2, 5, 6))
> mat[grep("P", rownames(mat)), 1] <- "#DCDCDC"
> mat[, 2] <- as.vector(mat[, 2])
> mat[which(is.na(mat[, 2])), 2] <- "#DCDCDC"
> par(las = 1, cex = 1.2, font = 2, mar = c(6, 6, 4, 2) + 0.1)
> dendromat(dhc, mat, edgePar = list(lwd = 2), nodePar = list(cex = 2, pch = "", font = 2)
```


1.2
1.0
0.8
0.6
0.4
0.2
0.0
response
subtype